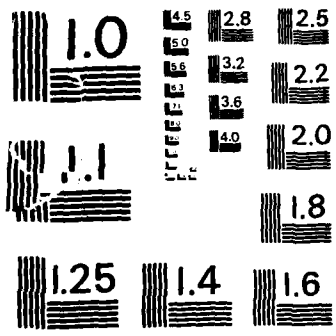
AD-A167 584 NONPARAMETRIC ESTIMATION OF QUANTILES AND OF DENSITY FUNCTIONS UNDER CENS... (U) SOUTH CAROLINA UNIV COLUMBIA DEPT OF MATHEMATICS AND STATISTI... W J PADGETT ET AL. 1/1

UNCLASSIFIED 14 JUN 85 AFOSR-TR-86-0174 AFOSR-84-0156 F/G 12/1 NL



UNCLASSIFIED

SECURITY CLASSIFICATION O

ON PAG

1a. REPORT SECURITY CLASS
UNCLASSIFIED

AD-A167 584

RICTIVE MARKINGS

2a. SECURITY CLASSIFICATIO

BUTION/AVAILABILITY OF REPORT

Approved for public release; distribution unlimited.

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Interim

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 86-0174

6a. NAME OF PERFORMING ORGANIZATION
Department of Mathematics and Statistics

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research

6c. ADDRESS (City, State and ZIP Code)
University of South Carolina
Columbia, SC 29208

7b. ADDRESS (City, State and ZIP Code)
Directorate of Mathematical & Information Sciences, Bolling AFB, DC 20332

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
AFOSR, ARO

8b. OFFICE SYMBOL (If applicable)
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-84-0156

8c. ADDRESS (City, State and ZIP Code)
Bolling AFB, DC 20332

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE (Include Security Classification)
ANNUAL TECHNICAL REPORT on Grant AFOSR-84-0156

12. PERSONAL AUTHOR(S)
W. J. Padgett and J. D. Spurrier

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Annual Technical Report	FROM 6/1/84 TO 5/31/85	1985 June 14	12

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Nonparametric quantile estimation; Density estimation; Right-censored data; Discrete failure models; Multiple comparisons; Accelerated life testing.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Major results have been obtained in the areas of nonparametric estimation of quantiles and of density functions under censoring, discrete failure models, and multiple comparisons. In particular, smooth nonparametric estimators of quantile functions from censored data were developed which give better estimates of percentiles of the lifetime distribution than the usual product-limit quantile function. Also, smooth density estimators from censored data were investigated using maximum penalized likelihood procedures. Several parametric models were proposed for the case of discrete failure data. These models provide a better fit to such data than some previously used discrete models. Finally, new methods of constructing simultaneous confidence intervals for pairwise differences of means of normal populations were developed, and the problem of selecting an asymptotically optimal design for comparing several new treatments with a control was solved. Work is continuing on the study of properties of kernel type quantile function estimators and development of goodness-of-fit tests for the model assumptions in accelerated life testing.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
MAJ Brian W. Woodruff

22b. TELEPHONE NUMBER (Include Area Code)
(202) 767-5027

22c. OFFICE SYMBOL
NM

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

1. Summary

This document reports the work performed and other research activities of the principal investigators during the one-year period from June 1, 1984, to May 31, 1985, under grant number AFOSR-84-0156.

In work under this grant, major results have been obtained in the areas of nonparametric estimation of quantiles and of density functions under censoring, discrete failure models, and multiple comparisons. In particular, smooth nonparametric estimators of quantile functions from censored data were developed which give better estimates of percentiles of the lifetime distribution than the usual product-limit quantile function. Also, smooth density estimators from censored samples were investigated using maximum penalized likelihood procedures. Several parametric models were proposed for the case of discrete failure data. These models provide a better fit to such data than some previously used discrete models. Finally, new methods of constructing simultaneous confidence intervals for pairwise differences of means of normal populations were developed, and the problem of selecting an asymptotically optimal design for comparing several new treatments with a control was solved. Work is continuing on the study of properties of kernel type quantile function estimators and development of goodness-of-fit tests for the model assumptions in accelerated life testing.

These results will provide useful methods which aid in the assessment of reliability and maintenance policies of various types of military equipment.

2. Research Objectives During the Reporting Period

The research objectives of this project can be divided into the following five categories:

(A) Nonparametric estimation of quantiles under censoring;

(B) Nonparametric density estimation from arbitrarily right-censored data;

(C) Discrete failure models;

(D) Goodness-of-fit tests for accelerated life testing models; and

(E) Multiple comparisons.

The specific research problems which were considered during this reporting period in these five areas will be outlined in this section. The progress toward their solution and significant results obtained to date will be described in Section 3 of this report.

A. Nonparametric Estimation of Quantiles Under Censoring

One of the main thrusts of this research was to develop smooth nonparametric estimates of the quantile function from arbitrarily right-censored observations. Such estimates using the kernel-smoothing method should yield better estimates of the percentiles of a lifetime distribution F^o than those obtained from the usual product-limit quantile function $\hat{Q}_n(p) = \inf\{t: \hat{F}_n(t) \geq t\}$, where $\hat{F}_n$ denotes the product limit estimator of F^o. Thus, median lifetime and other percentiles could be estimated more accurately, aiding in the development of maintenance and replacement policies for components or systems of interest.

B. Nonparametric Density Estimation from Arbitrarily Right-Censored Data

Another main area of investigation was to develop new estimators of probability density functions from right-censored data. One objective here

was to use the methods of maximum penalized likelihood estimation to obtain density estimates. This approach had not been studied under censoring and would result in smoother estimates than had been previously obtained from maximum likelihood methods without penalties for roughness. The first problem was to show that a maximum penalized likelihood estimator under censoring exists and is unique. Then the computation of the estimator was to be studied using numerical methods. Also, kernel-type density estimators under censoring were to be further studied with respect to optimal choice of bandwidth.

C. Discrete Failure Models

The objective in this area was to propose some new discrete parametric models for life testing situations in which the device fails after operating for a finite number of "cycles" or "time periods." These situations result in discrete data. Very few results for discrete failure models have appeared in the literature, and versatile models which would describe decreasing failure rates, increasing failure rates, or constant failure rates in the discrete data case were of interest. It was also of interest to develop estimation procedures for the parameters of the proposed models.

D. Goodness-of-Fit Tests for Accelerated Life Testing Models

Another objective of this research was to obtain powerful tests for simultaneously testing the model assumptions in accelerated life testing. In accelerated testing, it has been generally assumed that (i) the life distributions are the same except for possibly a scale change for the different stress levels and (ii) the scale parameters are related by a time transformation such as the inverse power law. Tests for these two assumptions and their performance characteristics were to be investigated.

E. Multiple Comparisons

In this area of investigation, methods for construction of exact simultaneous confidence intervals for pairwise differences of means of normal (lognormal) populations were to be studied. Also, the problem of selecting an appropriate optimal (in some sense) design for comparing several new treatments with a control was to be considered.

3. Status of the Research Effort

In this section a substantive statement of the progress and significant accomplishments towards achieving the research objectives outlined in Section 2 of this report will be given. The specific research papers containing the results will be listed in Section 4.

Professor Padgett is working primarily in the problem areas A and B listed in Section 2, and he and Professor Spurrier both worked on the investigation in area C. Professor Spurrier is performing the research associated with the areas D and E.

A. Nonparametric Estimation of Quantiles Under Censoring

Based on right-censored data from a lifetime distribution F^o, a kernel-type estimator of the quantile function $Q^o(p) = \inf\{t: F^o(t) \geq p\}$ for $0 \leq p \leq 1$ has been proposed in paper number [6] listed in Section 4. The estimator is defined by $\hat{Q}_n(p) = h_n^{-1}\int_0^1 \hat{Q}_n(t)K((t-p)/h_n)dt$, which is smoother than the usual product-limit quantile function $\hat{Q}_n(p) = \inf\{t: \hat{F}_n(t) \geq p\}$, where $\hat{F}_n$ denotes the product-limit estimator of F^o. Under the random censorship model and general (but nonrestrictive) conditions on h_n, K, and F^o, it is

shown that $Q_n(p)$ is strongly consistent. In particular, it is required for the strong consistency that K be a symmetric probability density with finite support which satisfies a Lipschitz condition and that F^o satisfy some continuity conditions. In addition, in paper [6] an approximation to Q_n is shown to be asymptotically equivalent to Q_n with probability one. In paper number [7] listed in Section 4, it has been shown that $Q_n(p)$ is asymptotically normal under conditions similar to those required for strong consistency. Also, Q_n and its approximation have been shown to be asymptotically equivalent in mean square under some slightly different assumptions on F^o, K, and h_n.

In paper number [6], a small Monte Carlo simulation study showed that for several values of the bandwidth h_n, $Q_n(p)$ performs better than $\hat{Q}_n(p)$ in the sense of smaller estimated mean squared errors. This is particularly true for the median lifetime estimate (i.e., $p = 0.5$). Further, a "data-based" choice of h_n was proposed by using the bootstrap. For a given censored sample, bootstrap estimates of the mean squared error of $Q_n(p)$ as a function of h_n were obtained. Then an estimate of h_n was obtained for computing $Q_n(p)$ from the given censored data by selecting the value, h_n^*, of h_n that gave small bootstrap estimated mean squared error of $Q_n(p)$.

The proposed nonparametric estimator Q_n is applicable to a broad spectrum of right-censoring situations, medical data as well as industrial life test data. Since no assumptions are required concerning the exact lifetime distribution, this estimator may be used in a variety of military equipment reliability and maintenance analyses.

The investigation of the properties of the smooth estimator Q_n and its approximation is still in progress. Work is being done currently on the mean convergence of the estimator and on relaxing the conditions on h_n for the asymptotic results.

B. Nonparametric Density Estimation from Arbitrarily Right-Censored Data

In paper number [1] listed in Section 4, the existence and uniqueness of the maximum penalized likelihood estimator (MPLE) of the density f^o of the lifetime distribution F^o is shown under the assumption of arbitrarily right-censored data. A penalty function ϕ is used to "penalize" rough estimators so that rough estimates are inadmissible. This is accomplished through the ϕ-penalized likelihood of the right-censored sample. An MPLE of f^o is then a function f which maximizes the ϕ-penalized likelihood subject to the condition that f be a probability density function. It was further shown in [1] that for Good and Gaskins' first penalty function the MPLE of f^o from right-censored data must be in the form of the solution to a linear integral equation whose forcing function is an exponential spline with knots at the observed censored and uncensored data points.

Current work in this area is concerned with the development of a numerical procedure to calculate the MPLE using the general form described above and some previous computation procedures for uncensored samples. The procedure will provide a smooth estimate of f^o from arbitrarily right-censored data using likelihood methods.

Also, in paper number [5], a summary of available nonparametric estimators of density and failure rate functions for right-censored samples is given. Included is a discussion of the choice of the bandwidth h_n based on likelihood for the given sample, i.e. a "data-based" procedure for choosing h_n.

C. Discrete Failure Models

A great amount of research has been performed for continuous lifetime distributions. However, very little has appeared in the literature for discrete

failure models. Discrete failure data arise in several situations: (i) A device can be monitored only once per time period (an hour, a day, etc.), and the observation taken is the number of time periods successfully completed prior to failure of the device. (ii) A piece of equipment may operate in cycles, and the experimenter observes the number of cycles successfully completed prior to failure.

Paper number [4] in Section 4 provides three families of discrete parametric lifetime distributions which are quite versatile in fitting increasing failure rate, decreasing failure rate, and constant failure rate models to either uncensored or right-censored life test data. The maximum likelihood estimation of the parameters, survival probabilities, and mean lifetimes has been investigated in that paper. The results of these studies indicate that the discrete models proposed will in some instances provide a better fit to discrete data than other discrete distributions given previously in the literature.

D. Goodness-of-Fit Tests for Accelerated Life Testing Models

During the reporting period, an investigation has begun into how to simultaneously test the life distribution assumption and the acceleration model assumption used in accelerated life testing. Accelerated life testing is a powerful technique for estimating the reliability of components that exhibit long lifetimes under normal operating conditions. However, these estimates are quite dependent upon the life and acceleration assumptions. Therefore, it is important to know if these assumptions are badly violated. While the investigation is not yet complete, the work to date is very encouraging. It appears likely that a simultaneous test of the life and acceleration

assumptions will be found that is more powerful than the existing test of only the life distribution. It is anticipated that this investigation will be completed during the next six months.

E. Multiple Comparisons

During this reporting period, two works have been completed in the area of multiple comparisons. Spurrier & Isham (1985) (paper number [2] in Section 4) present a method for construction of exact simultaneous confidence intervals for the pairwise differences of means of three normal populations with equal variances. This technique generalizes the well-known Tukey method to the case of unequal sample sizes. The computational technique involves writing the pivotal quantities in terms of a bivariate-t distribution. Numerical quadrature and Newton's method are used to find the exact probability point. The paper includes a table of probability points for small sample sizes. For larger sample sizes the probability points are found through an approximation based on the bivariate normal and studentized range distributions.

Spurrier & Edwards (1985) (paper number [3]) investigate the problem of selecting an incomplete block design for comparing several new treatments with a control or a standard treatment. A class of incomplete block designs, known as balanced treatment incomplete block designs, has been studied by several authors for use in this problem. Spurrier & Edwards show that a relatively small subclass of these designs contains the asymptotically optimal designs under the criterion of minimizing the length of simultaneous confidence intervals for treatment-control differences. A table of asymptotically optimal designs is presented. These results can be used to easily construct good designs in the case where one is using a large number of blocks.

4. Cumulative List of Written Publications in Technical Journals

In this section, the research papers that have been written under this grant are listed. They are divided into four categories: In print, accepted for publication, submitted, and research in progress. Copies of the manuscripts have been forwarded to the Program Manager as they were submitted for publication. Reprints of those already published will be forwarded as they become available.

A. In Print

[1] André M. Lubecke and W. J. Padgett, Nonparametric maximum penalized likelihood estimation of a density from arbitrarily right-censored observations. Communications in Statistics-Theory & Methods 14 (1985), 257-271.

[2] J. D. Spurrier and S. P. Isham, Exact simultaneous confidence intervals for pairwise comparisons of three normal means. Journal of the American Statistical Association 80 (1985), 438-442.

B. Accepted for Publication

[3] J. D. Spurrier and D. Edwards, An asymptotically optimal subclass of balanced treatment incomplete block designs for comparisons with a control. Biometrika (to appear).

[4] W. J. Padgett and J. D. Spurrier, On discrete failure models. IEEE Transactions on Reliability (to appear).

[5] W. J. Padgett, Nonparametric estimation of density and hazard rate functions when samples are censored. Handbook of Statistics, vol. 7 (to appear - invited).

C. Submitted

[6] W. J. Padgett, A kernel type estimator of a quantile function from right-censored data. Submitted to Journal of American Statistical Association.

[7] Y. L. Lio, W. J. Padgett, and K. F. Yu, On the asymptotic properties of a kernel type quantile estimator from censored samples. Submitted to Journal of Statistical Planning and Inference.

D. Research in Progress

[8] J. D. Spurrier, Goodness-of-fit tests for accelerated life testing.

[9] Y. L. Lio and W. J. Padgett, Further properties of kernel-type quantile estimators under random censorship.

5. Professional Personnel Associated with the Research Effort

In addition to the principal investigators, Professors Padgett and Spurrier, two graduate students have been supported (each partially) during this reporting period. André M. Lubecke was a research assistant during the summer, 1984. She is a doctoral student in statistics, and is currently completing her dissertation on the topic of maximum penalized likelihood estimation of density functions. Y. L. Lio is also a doctoral student in statistics who was supported as a research assistant for this grant during the academic year 1984-85. He has successfully completed the Admission to Candidacy Examination in the doctoral program.

6. Interactions

The principal investigators attended several professional meetings and conferences and gave (invited and contributed) talks as follows:

i) W. J. Padgett, J. D. Spurrier attended the Conference on Reliability and Quality Control, University of Missouri, Columbia, MO, June 4-8, 1984.

ii) W. J. Padgett (joint with A. M. Lubecke), "Nonparametric maximum penalized likelihood estimation of a density from arbitrarily right-censored data." Annual Meeting of IMS in Tahoe City, CA, August 21-24, 1984.

iii) J. D. Spurrier, "The training of statistical consultants at the University of South Carolina" (invited). Joint Statistical Meetings, Philadelphia, PA, August, 1984.

iv) D. Edwards, J. D. Spurrier, "An asymptotically optimal subclass of BTIB designs for comparisons with a control." Joint Statistical Meetings, Philadelphia, PA, August, 1984.

v) J. D. Spurrier and W. J. Padgett, "On discrete failure models." Spring Regional Meeting of the Biometric Society (ENAR) and ASA in Raleigh, NC, March 25-27, 1985.

vi) W. J. Padgett was the invited discussant for the Session on Reliability and Life Testing, Spring Regional Meeting of the Biometric Society (ENAR) and ASA in Raleigh, NC, March 25-27, 1985.

vii) W. J. Padgett, "Some estimation procedures for quantile and density functions from censored data." AFOSR Workshop on Reliability, Luray, VA, May 29-31, 1985.

In addition, Padgett gave a colloquium talk at Virginia Commonwealth University, Richmond, VA, and Spurrier gave colloquium talks at Shell Oil Research Center, Houston, TX, and at Oklahoma State University.

7. Inventions, Patent Disclosures, and Applications Stemming from the Research Project

No inventions or patents have stemmed from this research.

The results reported in Section 4 have wide application in the estimation and assessment of reliability and maintenance of military equipment. The various nonparametric procedures with censored data developed in this project allow accurate estimation of median lifetime, percentiles, densities, survival probability, and other lifetime characteristics without assuming particular forms of the life distribution. In particular, the kernel type quantile estimator will provide more accurate estimates of the median lifetime of a piece of equipment.

8. Other Professional Activities

During this reporting period, the principal investigators have been involved in numerous professional activities that are intimately related to the research efforts on this grant. Professors Padgett and Spurrier have refereed a total of fifteen manuscripts during the one-year period June 1, 1984, to May 31, 1985, one for Operations Research, three for Journal of the American Statistical Association, four for Communications in Statistics, three for Technometrics, two for IEEE Transactions on Reliability,one for Journal of Statistical Planning and Inference, and one for Journal of Multivariate Analysis. In addition, five reviews of papers were written for the Zentralblatt für Mathematik and one for Mathematical Reviews. Also, one proposal for the AFOSR was reviewed.

Professor Padgett was an Associate Editor of the Journal of Statistical Computation and Simulation and was a member of the Editorial Board of Communications in Statistics-Theory & Methods. Professor Spurrier was an Editorial Board Member of Communications in Statistics-Statistical Reviews.

Professor Padgett has served on the Program Committee for the 1985 SREB-ASA Summer Research Conference in Statistics, Boone, N. C., June 16-21, and is Program Co-Chairman for the 1986 Conference to be held in Mobile, Alabama, June 15-20. Also, Padgett chaired a session at the Conference on Reliability and Quality Control, University of Missouri, in June, 1984. In addition, Professor Spurrier organized a special seminar on "Quality Control" during the fall semester, 1984, at the University of South Carolina.

END

DTIC

6-86